THE AUTOMORPHIC TRANSFORMATION OF A BILINEAR FORM.

By J. H. M. Wedderburn.

1. **Introduction.** The problem of transforming a bilinear form into itself cogrediently was first solved by Hermite* and Cayley for symmetric and skew-symmetric forms and later by Voss† for any form. This solution has two defects; in the first place it only gives transformations whose determinant is $+1$, and secondly it becomes indeterminate for those transformations whose characteristic roots include both $+1$ and -1. These exceptional cases have been treated more or less completely by a number of authors.

The aim of this note is to present a method of obtaining a form for the automorphic transformation which displays clearly the rôle played by the exceptional cases. The parameters in this solution enter transcendentally but it is free from the first kind of exception; and in deriving from it the Hermite-Cayley form, which is rational, the analytical nature of this exceptional case is made clear.

The exponential and logarithmic functions of a matrix form the basis of the exposition and in view of this it has been thought advisable to include a short discussion of functions of a matrix in general especially as, in spite of the fact that there is little or nothing new in the results obtained, there is no place in the literature where the necessary properties are collected together.

2. **The idempotent units of a matrix.** If the elementary divisors of a matrix x are $(\lambda - g_i)^{p_i}$ $(i = 1, 2, \cdots, r)$, then, when the basis is properly chosen, x can be expressed‡ as the direct sum of irreducible matrices of the form

$$(1) \qquad x_i = \left\| \begin{matrix} g_i & 1 & & & \\ & g_i & 1 & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & g_i & 1 \\ & & & & g_i \end{matrix} \right\| = g_i e_i + \eta_i, \quad (i = 1, 2, \cdots, r)$$

* Hermite, "Sur la théorie des formes quadratiques ternaires indéfinies," Crelle, 47 (1854), pp. 307–312; Cayley, "A memoir on the automorphic linear transformation of a bipartite quadric," Lond. Phil. Trans., 148 (1858), pp. 39–46. For further references see Encyc. des Sci. Math., I, 2, fasc. 4, p. 489.

† Voss, "Über die cogredienten Transformationen einer bilinearen Form in sich selbst," Münch. Abh., 17 (1892), pp. 235–356.

‡ Cf. Bôcher, Higher algebra, p. 289.

where e_i and η_i are matrices of rank p_i and $p_i - 1$, respectively, which satisfy the conditions

(2) $e_i^2 = e_i, \quad \eta_i^{p_i} = 0, \quad \eta_i^{p_i-1} \neq 0, \quad e_i\eta_i = \eta_i = \eta_i e_i, \quad e_i e_j = 0 \quad (i \neq j).$

We shall say that e_i and η_i are the idempotent and nilpotent units corresponding to the elementary divisor $(\lambda - g_i)^{p_i}$. These units are not unique when the same root occurs in several elementary divisors. For instance, if x is the matrix

$$\left\|\begin{matrix} g & 0 & 0 \\ 0 & g & 1 \\ 0 & 0 & g \end{matrix}\right\|,$$

or, using matric units* e_{pq},

$$x = ge_{11} + g(e_{22} + e_{33}) + e_{23},$$

where $e_1 = e_{11}$, $e_2 = e_{22} + e_{33}$ and $\eta_2 = e_{23}$; then if we set

$$\begin{array}{lll} a_{11} = e_{11} - e_{13} & a_{12} = e_{12} & a_{13} = e_{13} \\ a_{21} = e_{21} - e_{23} & a_{22} = e_{22} & a_{23} = e_{23} \\ a_{31} = e_{11} - e_{13} + e_{31} - e_{33} & a_{32} = e_{32} + e_{12} & a_{33} = e_{33} + e_{13}, \end{array}$$

the a's form a set of matric units and

$$\begin{aligned} x &= g(e_{11} - e_{13}) + g(e_{22} + e_{33} + e_{13}) + e_{23} \\ &= ga_{11} + g(a_{22} + a_{33}) + a_{23}, \end{aligned}$$

so that $e_{11} - e_{13}$ and $e_{22} + e_{33} + e_{13}$ might have been chosen as idempotent units in place of e_1 and e_2.

It is shown below that, in any representation of x in the form (1), the sum e_j of all the idempotent units which belong to the same root g_j is independent of the particular representation used. It will be called the *principal* unit corresponding to g_j, its parts being called *partial* units.

It should be noticed that other normal forms are possible: for instance, in place of (1) we can by a different choice of η_i express x_i in the form

(3) $$x_i = g_i e_i + h_{i1}\eta_i + h_{i2}\eta_i^2 + \cdots + h_{i,\,p_i-1}\eta_i^{p_i-1},$$

where the h's are preassigned constants different from zero.

3. **Functions of a matrix.** Using the notation of the last paragraph, let

(4) $$x = \Sigma x_i, \qquad x_i = g_i e_i + \eta_i, \qquad (i = 1, 2, \cdots, r)$$

be an expression of x as the sum of irreducible matrices, then

$$x_i^m = (g_i e_i + \eta_i)^m = g_i^m e_i + mg_i^{m-1}\eta_i + \binom{m}{2} g_i^{m-2}\eta_i^2 + \cdots,$$

* The unit e_{pq} is a matrix for which the coefficient in the pth row and qth column is 1, while all the other coefficients are zero. The law of combination of these units is $e_{pq}e_{qr} = e_{pr}$, $e_{pq}e_{sr} = 0$ $(q \neq s)$.

the binomial expansion terminating with the $(m+1)$th term when m is less than the rank p_i of e_i, and with the p_ith term when this is not the case.

If now $f(\lambda)$ is any function expansible in a Taylor series which converges for every root of x, then $f(x)$ is reducible in the same way as x, the part corresponding to x_i being

$$(5)\quad f_i(x) = f(g_i)e_i + f'(g_i)\eta_i + f''(g_i)\frac{\eta_i^2}{2!} + \cdots + f^{(p_i-1)}(g_i)\frac{\eta_i^{p_i-1}}{(p_i-1)!}$$

or, writing (5) in full but with the subscript i omitted, we have

$$\left\|\begin{array}{cccccc} f(g) & f'(g) & \frac{f''(g)}{2!} & \frac{f'''(g)}{3!} & \cdots & \frac{f^{(p-1)}(g)}{(p-1)!} \\ & f(g) & f'(g) & \frac{f''(g)}{2!} & \cdots & \frac{f^{(p-2)}(g)}{(p-2)!} \\ & & f(g) & f'(g) & \cdots & \frac{f^{(p-3)}(g)}{(p-3)!} \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & \cdot \\ & & & & & f(g) \end{array}\right\|$$

where there are p_i rows and columns and every term to the left of the main diagonal is zero while, in the main diagonal itself and on its right, the terms in the first row are repeated in the succeeding rows, all terms lying on a parallel to the main diagonal being the same. An important particular case is $f(\lambda) \equiv \exp \lambda = \epsilon^\lambda$ for which

$$(6)\quad f_i(x) = \epsilon^{g_i}\left(e_i + \eta_i + \frac{\eta_i^2}{2!} + \cdots + \frac{\eta_i^{p_i-1}}{(p_i-1)!}\right) = e_i + x_i + \frac{x_i^2}{2!} + \frac{x_i^3}{3!} + \cdots.$$

Suppose now that $y = \Sigma y_i$ is a matrix whose reduced parts have the form

$$y_i = g_i^{(0)}e_i + g_1^{(1)}\eta_i + \cdots + g_i^{(p_i-1)}\frac{\eta_i^{p_i-1}}{(p_i-1)!},$$

where e_i and η_i are the same as in (4), i.e., belong to x, and the g's are any set of constants. Then by the extension of Lagrange's interpolation formula (see § 4 below), there is a polynomial $f(\lambda)$ for which

$f(g_i) = g_i^{(0)}, \quad f'(g_i) = g_i^{(1)}, \quad \cdots, \quad f^{(p_i-1)}(g_i) = g_i^{(p_i-1)}, \quad (i = 1, 2, \cdots)$

so that we may set $y = f(x)$. In particular if we set

$$g_i^{(0)} = \log g_i + 2k_i\pi\sqrt{-1} \equiv f(g_i)$$
$$g_1^{(j)} = \frac{d^j \log g_i}{dg_i^j} = f^{(j)}(g_i)$$

and also put $F(\lambda) = \epsilon^{f(\lambda)}$, then

$$F_i(x) = \epsilon^{f(g_i)}[e_i + f'(g_i)\eta_i + (f'^2(g_i) + f''(g_i))\eta_i^2 + \cdots]$$
$$= g_i e_i + \eta_i = x_i,$$

since the coefficients of the various powers of η_i are formally the successive derivatives of $\epsilon^{\log g_i}$. We have therefore $F(x) = x$, so that we may set $f(x) = \log x$. The logarithmic function so defined is indeterminate to an additive term of the form $2\pi i \Sigma k_i e_i$ where e_i $(i = 1, 2, \cdots)$ is any set of idempotent units belonging to x and the k's are integers. It is fairly obvious that any function which possesses the necessary derivatives may be extended to the case of a matric variable in a similar fashion.

Considerable care must be exercised in using the logarithmic function. For instance, if x and y are commutative, $\log x$ and $\log y$ will also be commutative if the same determination of $\log g_i$ is used with all the partial units depending on the root g_i; for the principal units of commutative matrices are commutative. If however this precaution is not taken, it is no longer true that $\log x$ and $\log y$ are necessarily commutative. For instance, if x is the matrix already used as an illustration in § 2 and $\operatorname{Log} g$ is a particular determination of $\log g$, then

$$z_1 = (\operatorname{Log} g + 2\pi i)e_{11} + \operatorname{Log} g\,(e_{22} + e_{33}) + \frac{1}{g}e_{23}$$

and

$$z_2 = \operatorname{Log} g\,(e_{11} - e_{13}) + \operatorname{Log} g\,(e_{22} + e_{33} + e_{13}) + \frac{1}{g}e_{23} = z_1 - 2\pi i e_{11}$$

are two determinations of $\log x$ which are not commutative. In this paper we shall only require logarithms in which the condition given above is satisfied and to indicate this we shall write $\operatorname{Log} x$ in place of $\log x$, so that $\operatorname{Log} x$ is determinate to an additive term of the form $\Sigma 2\pi k i e_j$ where the e_j are the *principal* units of x. The principal idempotent units of $\operatorname{Log} x$ are then the same as those of x while, as in (5), its principal nilpotent units are scalar polynomials of the corresponding principal nilpotent units of x.

The same difficulties arise, of course, with any multiple-valued function.

4. **The interpolation formula.** As we shall have need of it later, we shall now develop the generalization* of the Lagrange interpolation formula referred to in the previous section. Let

$$(7) \qquad \varphi(x) = (x - g_1)^{p_1}(x - g_2)^{p_2} \cdots (x - g_r)^{p_r}, \quad \left(\sum_1^r p_i = n, \;\; r > 1\right),$$

be the reduced equation of x, the roots g being all distinct. If we set

* Cf. Encyc. des Sci. Math., I, 2, fasc. 1, p. 61.

$$P_i(x) = \prod_{j \neq i} \left(\frac{x - g_j}{g_i - g_j}\right)^{p_j},$$

we can determine two polynomials $Q_i(x)$ and $D_i(x)$ of degree $p_i - 1$ and $n - p_i - 1$ respectively such that

$$P_i(x)Q_i(x) + (x - g_i)^{p_i}D_i(x) \equiv 1.$$

Setting

(9) $$R_i(x) = P_i(x)Q_i(x),$$

$1 - \Sigma_i R_i(x)$ is divisible by $\varphi(x)$ and, being of degree $n - 1$, at most is therefore zero; hence

(10) $$\sum_i R_i(x) \equiv 1.$$

If $h(x)$ is any polynomial in x with scalar coefficients, then

$$\begin{aligned} h(x) &= \Sigma h(x)R_i(x) \\ &= \Sigma\left[h(g_i) + h'(g_i)(x - g_i) + \cdots + \frac{h^{(p_i-1)}(g_i)}{(p_i - 1)!}(x - g_i)^{p_i-1}\right]R_i(x) \\ &\qquad + \Sigma C_i(x)(x - g_i)^{p_i}R_i(x), \end{aligned}$$

where C_i is a polynomial, being in fact the coefficient of $(x - g_i)^{p_i}$ in the remainder when $h(x)$ is expanded in a Taylor series. Now it follows from the definition of R_i that $(x - g_i)^{p_i}R_i(x)$ is divisible by $\varphi(x)$, hence, setting $R_{ij}(x) = (x - g_i)^j R_i(x)$,

(11) $$h(x) \equiv \sum_{i=1}^{r} \sum_{j=0}^{p_i-1} \frac{h^j(g_i)}{j!} R_{ij}(x) \pmod{\varphi(x)}.$$

If h is of lower degree than φ, this congruence is an algebraic identity, and therefore gives the form of a polynomial which, along with its derivatives up to the $(p_i - 1)$th order, has arbitrarily assigned values for $x = g_i$ $(i = 1, 2, \cdots, r)$.

When x is a matrix and $\varphi(x) = 0$ is its reduced equation, then (11) is again an identity in the *coefficients* of x and gives the form of any scalar polynomial in x. Since $R_i^2 = R_i$, $R_iR_j = 0$, $(i \neq j)$ and

$$R_{ij} = (x - g_i)^j R_i(x),$$

it is easily seen that R_i is what we have already called the principal idempotent unit belonging to g_i, and R_{i1} is the sum of the units η which belong to this root, i.e., it is the corresponding principal nilpotent unit; we may also notice here that $R_{i1}{}^j = R_{ij}$.

The principal units are therefore scalar polynomials in x, a result which is of some importance in the sequel.

The above argument requires some modification when h is not a

polynomial but, in view of what has already been said in the previous paragraph, it is not necessary to discuss the matter here.

Functions of two or more commutative matrices can be treated in a similar fashion.* Let x and y be two commutative matrices whose roots are $g_1, g_2, \cdots$ and $h_1, h_2, \cdots$, respectively, and as above let $R_i(x)$ and $R_i(y)$ $(i = 1, 2, \cdots)$ denote the principal units of these matrices. Then, if we set

$$S_{ij} = R_i(x)R_j(y),$$

those S_{ij} which are not zero are linearly independent. For, if $\Sigma\xi_{ij}S_{ij} = 0$, then

$$0 = R_p(x)\Sigma\xi_{ij}S_{ij}\cdot R_q(y) = \xi_{pq}S_{pq},$$

so that $\xi_{pq} = 0$ unless $S_{pq} = 0$.

From the definition of S_{ij} it follows that $S_{ij}S_{pq} = 0$ if $i \neq p$ or $j \neq q$, also $S_{ij}^2 = S_{ij}$ and $\Sigma_{ij}S_{ij} = 1$; hence

$$x = \sum_{i,j}[g_i + (x - g_i)]S_{ij}, \qquad y = \sum_{i,j}[h_j + (y - h_j)]S_{ij},$$

where $(x - g_i)S_{ij}$ and $(y - h_j)S_{ij}$ are commutative nilpotent matrices.

If $\psi(x, y)$ is any scalar polynomial in x and y, we may now set

$$\begin{aligned}\psi(x, y) &= \sum_{r,s}\psi_{rs}{}^{ij}(x - g_i)^r(y - h_j)^s \\ &= \sum_{i,j}\left[\psi(g_i, h_j)S_{ij} + \sum_{r,s}\psi_{rs}{}^{ij}(x - g_i)^r(y - h_j)^sS_{ij}\right],\end{aligned}$$

where in the second summation r and s are not both zero; or, if we let $(x - g_i)^r(y - h_j)^sS_{ij} = S_{ij}{}^{rs}$, then

$$\begin{aligned}\psi(x, y) &= \sum_{i,j}\psi(g_i, h_j)S_{ij} + \sum_{i,j}\sum_{r,s}\psi_{rs}{}^{ij}S_{ij}{}^{rs} \\ &= z + w,\end{aligned}$$

say, where w is nilpotent, being the sum of a number of commutative nilpotent matrices. Now if $\varphi(z) = 0$ is the reduced equation of a matrix z, and w is a nilpotent matrix commutative with z for which $w^s = 0$, then, if $F(z) = \varphi^s(z)$, we have

$$F(z + w) = F(z) + F'(z)w + \cdots + \frac{F^{(s-1)}(z)}{(s-1)!}w^{s-1} = 0,$$

since the first s derivatives of $F(z)$ are divisible by $\varphi(z)$ and therefore vanish. It follows that the characteristic of the reduced equation of $z + w$ is a factor of a power of that of z, and vice versa; hence the roots of z and $z + w$ are the same. We can say, therefore, that *if $R_i(x)$ and $R_j(y)$ $(i, j = 1, 2, \cdots)$ are the principal idempotent units of two commuta-*

* Cf. Frobenius, "Über vertauschbare Matrizen," Berl. Sitzb. (1896), pp. 601–614.

tive matrices x and y, and $S_{ij} = R_i(x)R_j(y)$; and if g_i and h_j are the corresponding roots of x and y; then the roots of any scalar function $\psi(x, y)$ of x and y are $\psi(g_i, h_j)$ where i and j take only those values for which $S_{ij} \neq 0$.

The extension to functions of several commutative matrices is obvious.

5. **The automorphic transformation of a matrix.** If y is a non-singular matrix, the problem of transforming it into itself is equivalent to finding all the matric solutions of the equation*

$$x'yx = y. \tag{12}$$

When solved for x', this equation gives

$$x' = yx^{-1}y^{-1}, \tag{13}$$

from which it follows immediately that the identical equation of x has reciprocal roots and that, if g is any root other than ± 1, the elementary divisors corresponding to g and $1/g$ occur in pairs† with the same exponents. It follows also from (13) that

$$x' = yx^{-1}y^{-1} = y'x^{-1}y'^{-1}$$

so that x is commutative with $y^{-1}y'$.

If $h(\lambda)$ is a scalar polynomial in λ, then from (13) $h(x') = yh(x^{-1})y^{-1}$, and therefore, in particular, the principal unit of x' corresponding to a root $g_i \neq \pm 1$ is the transform of the principal unit‡ of x corresponding to $1/g_i$. If we denote the principal units belonging to g_i ($g_i \neq \pm 1$) and $1/g_i$ by e_i and e_{-i}, respectively, we have therefore

$$e_i' = ye_{-i}y^{-1}; \tag{14}$$

and similarly, if e_1 and e_{-1} belong to the roots ± 1 when these roots are present, we have

$$e_1' = ye_1y^{-1}, \qquad e_{-1}' = ye_{-1}y^{-1}. \tag{15}$$

If now we set

$$x = \epsilon^z, \qquad z = \operatorname{Log} x, \tag{16}$$

then from (12)

$$1 = x'yxy^{-1} = \epsilon^{z'}\epsilon^{yzy^{-1}} = \epsilon^{z'+yzy^{-1}}. \tag{17}$$

Here $z' + yzy^{-1}$ has the same principal idempotent units as x' and z', and hence it has the form $\Sigma(\gamma_i e_i' + \bar{\eta}_i')$ where $\bar{\eta}_i'$ is a scalar polynomial in η_i'; hence (17) is equivalent to

$$1 = \epsilon^{z'+yzy^{-1}} = \Sigma\epsilon^{\gamma_i}\left(e_i' + \bar{\eta}_i' + \frac{\bar{\eta}_i'^2}{2!} + \cdots\right),$$

* Here x' denotes, as usual, the transverse or conjugate of x.

† Cf. Kronecker, Crelle, 68 (1868), p. 273.

‡ Cf. Taber, "On the automorphic linear transformation of an alternate bilinear form," Math. Ann., 46 (1895), p. 568. The principal unit of x belonging to g is the same as the principal unit of $1/x$ belonging to $1/g$.

whence $\bar{\eta}_i' = 0$ and* $\gamma_i = 2k_i\pi\iota$. We can therefore set

$$(18) \qquad z' + yzy^{-1} = 2\pi\iota\Sigma k_i e_i'.$$

Since x, and therefore also z, is commutative with $y^{-1}y'$, we have

$$y'^{-1}z'y' + z = 2\pi\iota\Sigma k_i y'^{-1}e_i'y'$$

or, forming the transverse of each side and using (14) and (15),

$$z' + yzy^{-1} = 2\pi\iota\Sigma k_i ye_i y^{-1} = 2\pi\iota(k_1e_1' + k_{-1}e_{-1}' + \Sigma k_i e_i') \quad (i \neq \pm 1),$$

and, comparing this with (18), we have

$$(19) \qquad k_i = k_{-i} \qquad (i \neq \pm 1).$$

We can now simplify (18) as follows. Set

$$z_1 = z - 2\pi\iota(\Sigma' k_i e_i + \lambda_1 e_1 + \lambda_{-1}e_{-1}),$$

where in the summation sign the prime indicates that the roots $g_i \neq \pm 1$ are arranged in pairs g_i and $1/g_i$ and only the first of each pair is taken in forming the sum. Inserting z_1 in place of z in (18) we have from (14) and (15)

$$\begin{aligned} z_1' + yz_1y^{-1} &= z' + yzy^{-1} - 2\pi\iota(\Sigma' k_i e_i' + \lambda_1 e_1' + \lambda_{-1}e_{-1}' \\ &\qquad + \Sigma' k_i e_{-i}' + \lambda_1 e_1' + \lambda_{-1}e_{-1}') \\ &= 2\pi\iota[(k_1 - 2\lambda_1)e_1' + (k_{-1} - 2\lambda_{-1})e_{-1}'], \end{aligned}$$

where by a proper choice of λ_1 and λ_{-1} the coefficients of e_1' and e_{-1}' may be made equal to 0 or 1. Now evidently $\epsilon^{z_1} = \epsilon^z$; it follows that there is no lack of generality in writing in place of (18)

$$(20) \qquad z' + yzy^{-1} = 2\pi\iota\zeta',$$

where

$$(21) \qquad \zeta = \theta_1 e_1 + \theta_2 e_{-1} = \zeta_1 + \zeta_2 \qquad (\theta_1, \theta_2 = 0 \text{ or } 1).$$

Writing now

$$(22) \qquad w = z + \pi\iota\zeta,$$

equation (20) becomes

$$(23) \qquad w' + ywy^{-1} = 0,$$

which may also be written

$$y'(wy^{-1})' + y(wy^{-1}) = 0,$$

or, if $u = wy^{-1}$,

$$(23') \qquad y'u' + yu = 0,$$

which is equivalent to the equation given by Cayley.† The solution of

* Here $\iota = \sqrt{-1}$.

† Cayley, l.c., p. 44. Cayley's solution is incomplete as he omits to impose the necessary conditions on the skew-symmetric matrix which enters into his result; and this leads him to draw erroneous conclusions.

(23) which is given in the next section is practically that given by Voss.*

6. **The equation** $w' + ywy^{-1} = 0$. We shall consider in place of (23) the more general equation

$$w' = \delta ywy^{-1} \qquad (\delta = \pm 1). \tag{24}$$

Forming the transverse of each side, we get $w = \delta y'^{-1}w'y'$ or $w' = \delta y'wy'^{-1}$, whence

$$wy^{-1}y' = y^{-1}y'w \qquad \text{or} \qquad y'wy'^{-1} = ywy^{-1}, \tag{25}$$

i.e., w is commutative with $y^{-1}y'$. Now from (24) we have $w = \delta y^{-1}w'y$, so that $2w = w + \delta y^{-1}w'y$. But if v is any matrix commutative with $y^{-1}y'$, then

$$w = v + \delta y^{-1}v'y \tag{26}$$

is a solution of (24) as, on substituting this value for w, we get

$$w' - \delta ywy^{-1} = v' + \delta y'vy'^{-1} - \delta yvy^{-1} - v' = 0,$$

since $y'vy'^{-1} = yvy^{-1}$. The most general solution of (23) is therefore obtained by setting

$$w = v - y^{-1}v'y, \qquad vy^{-1}y' = y^{-1}y'v. \tag{27}$$

It should be noted, however, that two different values of v may lead to the same value of w.

When $\delta = -1$, we have relations among the roots and idempotent units of w which are the logarithmic counterpart of those already given for x. For, since

$$|\lambda - w| = |\lambda - w'| = |\lambda + ywy^{-1}| = |\lambda + w|, \tag{28}$$

the non-zero roots of w occur in pairs of opposite sign and with equal exponents in the elementary divisors. We can show exactly as in § 5 that if e_i is the principal unit corresponding to a root g_i ($g_i \neq 0$) and e_{-i} the principal unit belonging to $-g_i$, then

$$e_i' = ye_{-i}y^{-1}, \qquad e_{-i}' = ye_iy^{-1}, \tag{29}$$

and if e_0 is the principal unit belonging to the root 0, if present, then

$$e_0' = ye_0y^{-1}. \tag{30}$$

Since $(w')^r = (-1)^r yw^ry^{-1}$, the reduced equation of w has the form $w^m\psi(w^2) = 0$; hence e_0 is a polynomial in w^2, which gives an independent proof of (30), since $(w^2)' = yw^2y^{-1}$.

The form of w given in (22) can be still further simplified by means of these relations. In (21) the term ζ_1 is the sum of partial units of x

* Voss, l.c., p. 330.

coming from roots equal to unity and they therefore correspond to roots of the form $2\pi k\iota$ of z, where k is integral, and hence to roots $\pi(2k+1)\iota$ of w. Let a_1 be the principal unit of w corresponding to this root and a_2 that belonging to its negative $-\pi(2k+1)\iota$ so that by (29)

$$a_1' = ya_2y^{-1};$$

and let ζ_{11} be that part of ζ_1 which is a partial unit of a_1wa_1 so that $\zeta_{11} = a_1\zeta_1 = \zeta_1 a_1$; then, if $\zeta_{22} = \zeta_1 a_2 \zeta_1$, we have

$$\zeta_{11}' = \zeta_1' a_1' \zeta_1' = y\zeta_1 a_2 \zeta_1 y^{-1} = y\zeta_{22}y^{-1};$$

and similarly

$$\zeta_{22}' = y\zeta_{11}y^{-1}.$$

But $\zeta_{11} = \zeta\zeta_{11}\zeta$; therefore $\zeta_{11}' = y\zeta\zeta_{22}\zeta y^{-1}$, so that $\zeta\zeta_{22}\zeta = \zeta_{22}$ and ζ_{22} is therefore also a partial unit along with ζ_{11}; the rank of $\zeta_1 - \zeta_{11} - \zeta_{22}$ is less than that of ζ_1.

If now we put

$$\bar{z} = z - 2\pi\iota\,\zeta_{11},$$

we have as before

$$\begin{aligned} \bar{z}' + y\bar{z}y^{-1} &= 2\pi\iota(\zeta_1' + \zeta_2') - 2\pi\iota\zeta_{11}' - 2\pi\iota y\zeta_{11}y^{-1} \\ &= 2\pi\iota(\zeta_1' - \zeta_{11}' - \zeta_{22}') + 2\pi\iota\zeta_2'. \end{aligned}$$

This transformation therefore replaces ζ_1 by a new ζ with lower rank and at the same time does not alter x. By repeating this process we can reduce the rank to zero which means that we can assume $\zeta_1 = 0$ without loss of generality.

In the same way ζ_{-1} corresponds to roots $(2k+1)\pi\iota$ of z and therefore to roots $(2k+2)\pi\iota$ of w. If $k \neq -1$, the rank of ζ_{-1} can be reduced as above so that it is only necessary to take account of zero roots of w in considering the form of ζ_{-1}.

7. **The determination of** z. The results of the preceding paragraph may be summarized by saying that every value of x in (12) can be obtained by putting

$$z = w + \theta\pi\iota\zeta, \qquad (\theta = 0, 1) \tag{31}$$

where w is any solution of (23) and ζ is an idempotent matrix corresponding to a zero root of w which satisfies the equation

$$\zeta' = y\zeta y^{-1}. \tag{32}$$

In order to complete the determination of z it is therefore necessary to show how ζ is to be determined. The principal idempotent unit, e_0, belonging to the zero root* of w is of course one possible value; the only

* When the order, n, of y is odd, there must evidently be an odd number of such roots; while if n is even, there will be an even number or none.

difficulty is then to ascertain when there will exist partial units of e_0 which satisfy (32).

We shall first separate off the part of w depending on e_0 by writing

$$w = (1 - e_0)w + e_0 w = w_1 + w_0,$$

where $w_1 w_0 = 0 = w_0 w_1$. The zero roots of w_1 correspond to simple elementary divisors, and w_0 has only zero roots; both are solutions of (23).

Suppose now that e_0 can be expressed as the sum of partial units of x, say

$$e_0 = e_1 + e_2 + \cdots + e_p,$$

such that $e_i e_j = 0$ $(i \neq j)$ and $e_i' = y e_i y^{-1}$; each e_i is then a possible determination of ζ. This being so, the matrix

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_p e_p,$$

where the α's are scalars, is a solution of

$$(33) \qquad a' = y a y^{-1},$$

$e_1, e_2, \cdots$ being the principal units corresponding to the roots $\alpha_1, \alpha_2, \cdots, \alpha_p$. Conversely, if a is any solution of (33), $e_0 a e_0$ is also a solution and if $e_1, e_2, \cdots, e_s$ are its principal units, they are solutions of (32) and are therefore available as values of ζ. Also if ζ is a sum of any or all of these e's, then $w_0 = (1 - \zeta)w_0 + \zeta w_0$, and $(1 - \zeta)w_0$ and ζw_0 are solutions of (23) all of whose roots are zero. Further, if e is any idempotent unit which satisfies (33), ewe is a solution of (23). We therefore conclude that *every x which transforms a non-singular matrix y into itself cogrediently is of the form ϵ^z where z is determined as follows: take any solution w_1 of* (23) *and any solution a of* (33) *and let ζ be a principal unit of the latter, then*

$$z = (1 - \zeta)w_1(1 - \zeta) + \zeta w_1 \zeta + \theta\pi\iota\zeta = w + \theta\pi\iota\zeta, \quad (\theta = 0, 1).$$

The determinant of x is ± 1 according as the rank of ζ is odd or even. Here w may be any solution of (23) and is therefore a continuous function of a certain number of parameters; hence x is also a continuous function of these parameters but involves at least one other parameter θ in which it is not continuous since the part of $\log x$ which depends on θ vanishes except for $\theta = 1$.

We may also notice here that we can set

$$x = \epsilon^w(1 - 2\theta\zeta),$$

where $(1 - 2\theta\zeta)^{-1} = (1 - 2\theta\zeta)$; and if $w = (1 - e_0)w(1 - e_0) + e_0 w e_0 = w_1 + w_0$, e_0 being the principal unit of w corresponding to its zero root,

then w_0 is nilpotent and, if

$$\gamma = w_0 + \frac{w_0^2}{2!} + \cdots = e^{w_0} - 1, \tag{34}$$

then

$$x = \epsilon^{w_1}(1 + \gamma)(1 - 2\theta\zeta), \qquad (\theta = 0, 1). \tag{35}$$

Here w_1 is any solution of (23) in which the zero roots have simple elementary divisors, e_0 is the principal idempotent unit corresponding to the zero root* of w_1, and γ and ζ are matrices which are respectively nilpotent and idempotent and are both solutions of

$$\varphi = e_0\varphi e_0, \qquad \varphi' = y\varphi y^{-1}.$$

8. **Rational parameters.** The parameters involved in the form of x given in the preceding paragraph enter transcendentally. If however we set

$$t = \tanh\frac{z}{2} = \frac{\epsilon^z - 1}{\epsilon^z + 1} = \frac{x - 1}{x + 1}, \tag{36}$$

then $\epsilon^z = (t - 1)/(t + 1)$ or

$$x = \frac{t - 1}{t + 1}, \tag{37}$$

also

$$yty^{-1} = \frac{yxy^{-1} - 1}{yxy^{-1} + 1} = \frac{x'^{-1} - 1}{x'^{-1} + 1} = \frac{1 - x'}{1 + x'} = -t',$$

so that t is a solution of (23); and if the coefficients of t are taken as parameters in so far as they are independent, (37) expresses x rationally in terms of these parameters. If, however, $|x + 1| = 0$, t becomes infinite, so that this form cannot give any solution which has roots equal to -1, at least directly. The difficulty arises from the fact that $\tanh(\theta/2) \to \infty$ as $\theta \to \pi\iota$, but, since $(t - 1)/(t + 1) \equiv \epsilon^z$ for all values of t which do not possess an infinite root, i.e., a root corresponding to a root $(2k + 1)\pi\iota$ of z, then x will be a solution of (12) so long as the coefficients of z are continuous functions of the parameters involved and the limiting value of x is finite and determinate. Now z is a continuous function of the parameters involved in v in equation (26) but is in general discontinuous in the coefficients of ζ; and moreover a ζ term is present only when x has a root -1. Hence if z is a solution of (17) which has no root equal to an odd multiple of $\pi\iota$, then t is finite and the expression for x in (37) remains finite even if t becomes infinite so long as z is finite and has no ζ term.

9. **Automorphic transformation of symmetric and skew-symmetric matrices: orthogonal matrices.** If y is symmetric or skew-symmetric, the matrix v

* If w_1 has no zero root, γ, θ and ζ are equal to 0.

occurring in the solution of (24) is entirely arbitrary since $y^{-1}y' = \pm 1$. Hence from (26)

$$w = v + \delta y^{-1}v'y = (vy^{-1} + \delta y^{-1}v')y = uy,$$

where $u = vy^{-1} + \delta y^{-1}v'$. Taking $\delta = -1$, u is skew-symmetric if y is symmetric and vice versa, and as any skew-symmetric (symmetric) matrix can be put in this form, u may be taken to be an arbitrary skew-symmetric (symmetric) matrix. Similarly if $\delta = +1$, the value of a in (33) becomes $a = by$, where b is an arbitrary symmetric (skew-symmetric) matrix.

If $y = 1$, then w is skew-symmetric and a symmetric; hence every orthogonal matrix has the form

$$x = \epsilon^{w}(1 - 2\theta\zeta), \qquad (\theta = 0, 1)$$

where ζ is a symmetric idempotent matrix (which may be zero) and w is a skew-symmetric matrix commutative with ζ. The known theorems regarding the roots of real orthogonal matrices are readily derived from this form.